Raindrop size distribution estimation from S-band polarimetric radar using a regularized neural network

Gianfranco Vulpiani\textsuperscript{1}, Frank S. Marzano\textsuperscript{2}, V. Chandrasekar\textsuperscript{3}, Alexis Berne\textsuperscript{4}, Remko Uijlenhoet\textsuperscript{5}

\textsuperscript{1} Meteo France, Trappes (France).
\textsuperscript{2} University La Sapienza, Rome (Italy).
\textsuperscript{3} Colorado State University, Fort Collins (Colorado).
\textsuperscript{4} LTHE, Grenoble (France).
\textsuperscript{5} Wageningen University, Wageningen (The Netherlands).

1 Introduction

Radar rainrate estimates are prone to a high degree of uncertainty due to several error sources, among which the space-time variability of the raindrop size distribution (RSD). Polarimetric radar measurements enable the use of combined algorithms which reduce the sensitivity to the variability of the RSD. The aim of this work is to develop a new procedure to retrieve the RSD parameters which can be used to estimate the corresponding rainfall rates. The polarimetric variables (\(Z_{hh}\), \(Z_{dr}\), and \(K_{dp}\)) are used to retrieve the RSD parameters by means of an ad-hoc neural network (NN) technique. The reason for this choice is the ambition to exploit the capability of NNs to approximate strongly non-linear functions such as those describing the relationships between radar observables and RSD parameters. A stochastic model, based on disdrometer measurements, is used to generate realistic range profiles of raindrop size distribution parameters while a T-matrix solution technique is adopted to compute the corresponding polarimetric variables at S band.

2 Polarimetric scattering model of rainfall

A gamma raindrop size distribution (RSD), having the general form \(N(D) = N_0 \cdot D^\alpha \cdot \exp(-A \cdot D)\) with \(D\) the particle diameter and \(N_0\), \(A\) and \(\mu\) RSD parameters, has been introduced in the literature to account for most of the variability occurring in the naturally observed RSD. The concept of normalization has been introduced by Willis (1984) and revisited by Chandrasekar and Bringi (1987), and Testud et al. (2001). The number of raindrops per unit volume per unit size can be written as:

\[
N(D) = N_0 \cdot f(\mu) \left( \frac{D}{D_0} \right)^\alpha \exp\left[ -(3.67 + \mu) \frac{D}{D_0} \right] \tag{1}
\]

where \(f(\mu)\) is a function \(\mu\) only, the parameter \(D_0\) is the median volume drop diameter, \(\mu\) is the shape parameter of the drop spectrum, and \(N_0\) [\text{mm}^{-1} \text{m}^{-3}] is a normalized drop concentration that can be calculated as function of liquid water content \(W\) and \(D_0\) (e.g., Bringi and Chandrasekar, 2001).

2.1 Polarimetric radar variables

The copolar radar reflectivity factors \(Z_{hh}\) and \(Z_{vv}\) [\text{mm}^6 \text{m}^{-3}] at H and V polarization state and the differential reflectivity \(Z_{dr}\) [\text{dB}] can be expressed as follows:

\[
Z_{hh,vv} = \frac{\lambda^2}{\pi^2 |K|^2} < 4\pi |S_{hh,vv}(D)|^2 > \tag{2}
\]

\[
Z_{dr} = 10 \log_{10} \left( \frac{Z_{hh}}{Z_{vv}} \right) \tag{3}
\]

where \(S_{hh,vv}\) are the backscattering co-polar components of the complex scattering matrix \(S\) of a raindrop and the angular brackets represent the ensemble average over the RSD. \(K\) depends on the complex dielectric constant of water estimated as a function of wavelength \(\lambda\) [\text{mm}] and temperature. For a polarimetric radar, the specific differential phase shift \(K_{dp}\) [\text{° km}^{-1}], due to the forward propagation phase difference between H and V polarization can be obtained as:

Correspondence to: Gianfranco Vulpiani.
gianfranco.vulpiani@meteo.fr
\[ K_{dp} = 10^{-3} \frac{180}{\pi} \lambda \cdot \text{Re}[\langle f_{lh}(D) - f_{rv}(D) \rangle] \quad (4) \]

where \( f_{lh,rv} \) are the forward-scattering co-polar components of S. The rainfall rate \( R \) [mm h\(^{-1}\)] is defined as
\[ R = 6 \pi 10^{-4} \int D^3 v(D) N(D) dD \quad (5) \]

where \( v(D) \) [m s\(^{-1}\)] is the raindrop terminal fall speed in still air. At sea level, a commonly used raindrop speed relationship is the one derived by Atlas and Ulbrich (1977).

### 3 Neural network retrieval technique

#### 3.1 Minimization and regularization techniques

An artificial neural network is a non-linear parameterized mapping from an input \( \mathbf{x} \) to an output \( y = \mathbf{NN}(\mathbf{x}; \mathbf{w}, \mathbf{M}) \), where \( \mathbf{w} \) is the vector of parameters relating the input to the output while the functional form of the mapping (i.e., the architecture of the net) is denoted as \( \mathbf{M} \). The multi-layer perceptron architecture (MLP), considered here, is a mapping model composed of several layers of parallel processors. It has been theoretically proven that one-hidden layer MLP networks may represent any non-linear continuous function (Haykin, 1995), while a two-hidden layer MLP may approximate any function to any degree of accuracy (Battiti, 1989) has been implemented in this work. Gradient descent may get stuck in local minima of the performance function. The best strategy in this case is to orient the search towards the global minimum, but the form of the error function may be such that the gradient does not point in this direction. The problem can be overcome by including a momentum term \( m \) in the weight updates (Faussett, 1994). In the present study a value of \( m = 0.9 \) has been assumed (Hagan, 1996). The ideal neural network is characterized by small errors on the training set and the capability to respond properly to new inputs. The latter property is called generalization. The procedure to improve generalization, called regularization, adds an additional term to the objective function which becomes \( E_{\text{R}} = E_{\text{N}} - (1 - \lambda) E_{\text{W}} \). \( E_{\text{W}} \) is the sum of squares of the networks weights and biases. In addition we note that, in Aires et al. (2002), the authors have experimentally proven that, for noisy data, a one-hidden layer MLP network may improve the network generalization through the reduction of the number of parameters. In Bishop (1996), the author demonstrated that, assuming low noise conditions, training with input perturbation is closely related to regularization.

#### 3.2 Raindrop size distribution retrieval

Reflectivity and differential reflectivity are commonly used for RSD retrieval (Gorgucci et al., 2002; Brandes et al., 2002). Specific differential phase shift is a potential predictor for RSD estimation but it may be affected by a high noise which may perturb the results. Consequently, \( K_{dp} \) has been used in Gorgucci et al. (2002) setting a lower threshold of 0.2 deg km\(^{-1}\). It has been found that the proposed algorithm performs well even for very low values of \( K_{dp} \). Moreover, in case of unreliable or unavailable measurements of \( K_{dp} \) (i.e., radars which do not measure differential phase), a 2-input neural network algorithm can also be successfully used. The median volume drop diameter \( D_v \) and the intercept parameter \( N_v \) are independently estimated using distinct NNs with 3 (i.e., \( Z_{hh}, Z_{dr}, K_{dp} \)) or 2 inputs (i.e., \( Z_{hh}, Z_{dr} \)) according to the availability and reliability of \( K_{dp} \). The shape parameter \( \mu \) is estimated from \( Z_{dp} \) and the retrieved values of \( D_v \) (as suggested in Brandes, 2002) using 2-input (i.e., \( Z_{dr}, D_v \)) NN. Consequently, its estimate is indirectly dependent on \( K_{dp} \) through \( D_v \). The proposed RSD retrieval technique can be formalized in the following way:

\[
\begin{align*}
\text{NN}_{D_0}(Z_{hh}, Z_{dr}, K_{dp}) \\
\text{NN}_{N_v}(Z_{hh}, Z_{dr}, K_{dp}) \\
\text{NN}_{\mu}(Z_{dr}, D_v)
\end{align*}
\]

if \( K_{dp} \) is available, \( \text{NN}_{K_{dp}}(Z_{hh}, Z_{dp}) \) and \( \text{NN}_{N_v}(Z_{hh}, Z_{dp}) \) being used otherwise. The neural network architecture and regularization parameters have been determined according to a heuristic monitoring of the generalization capability on test data, the root mean square error having been used as metric. According to what is suggested in Aires (2002), it has been sensitive to the proper setting of the learning rate. For this reason, a back-propagation training with an adaptive learning rate is crucial. Battiti’s “bold driver” technique (Battiti, 1989) has been implemented in this work. Gradient descent may get stuck in local minima of the performance function. The best strategy in this case is to orient the search towards the global minimum, but the form of the error function may be such that the gradient does not point in this direction. The problem can be overcome by including a momentum term \( m \) in the weight updates (Faussett, 1994). In the present study a value of \( m = 0.9 \) has been assumed (Hagan, 1996). The ideal neural network is characterized by small errors on the training set and the capability to respond properly to new inputs. The latter property is called generalization. The procedure to improve generalization, called regularization, adds an additional term to the objective function which becomes \( E_{\text{R}} = E_{\text{N}} - (1 - \lambda) E_{\text{W}} \). \( E_{\text{W}} \) is the sum of squares of the networks weights and biases. In addition we note that, in Aires et al. (2002), the authors have experimentally proven that, for noisy data, a one-hidden layer MLP network may improve the network generalization through the reduction of the number of parameters. In Bishop (1996), the author demonstrated that, assuming low noise conditions, training with input perturbation is closely related to regularization.
found that the one-hidden layer configuration improves the
generalization capability of the NNs. The number of nodes
in the hidden layer has been fixed to 6 for \( NN_{D0} \) and \( NN_{Nw} \),
and 12 for \( NN_{\mu} \). A choice of \( \gamma = 0.7 \) has been found to be
suitable for the retrieval of all the RSD parameters. The
input perturbation technique has also been adopted in order
to increase the generalization of \( NN_{\mu} (Z_{dr}, D_0) \).

Gaussian noise with standard deviation equal to 1 dBZ, 0.2
dB and 0.3 ° km\(^{-1}\) has been added to \( Z_{hh} \), \( Z_{dr} \) and \( K_{dp} \),
respectively.

The sensitivity analysis has been accomplished in terms of
mean error \( \epsilon \), root mean square error RMSE and correlation
coefficient \( r \). Figure 1 shows the comparison between the
proposed technique and those described in Gorgucci et al.
(2002) (named \( \beta \) method) and Brandes et al. (2002) (named
C-G method). The median volume drop diameter is
estimated fairly well by all examined algorithms. The
differences in terms of RMSE are small but the proposed
approach shows a larger correlation between the estimated
and true RSD parameters. We found that \( \epsilon(\hat{D}_0, D_0) \) is
0.941, 0.856 and 0.796, respectively, for the NN, C-G and
\( \beta \) algorithms. Moreover, the estimation of \( N_w \) by means of the
\( \beta \) algorithm is affected by a larger error standard deviation
and smaller correlation coefficient as compared to the other
examined methods. The latter have marked differences,
especially in terms of correlation, \( \epsilon(\hat{N}_w, N_w) \) being 0.943
and 0.854, respectively, for the NN and C-G methods. Once
the RSD parameters are estimated using the mentioned

4 Numerical Results on synthetic data

The choice of the architecture, learning algorithm and
performance function is crucial for the neural network set
up. On the other hand, the training data set also plays an
important role to guaranty a high generalization capability.
In order to “regularize” the neural network and test its
robustness, a large highly heterogeneous data set has been
generated by randomly varying the RSD parameters, the axis
ratio relationship and the temperature. The validation of the
proposed methodology has been accomplished using realistic
range profiles of RSDs generated using a modified version
of the stochastic simulator proposed by Berne and Uijlenhoet
(2005). It is worth mentioning that it is based on a gamma
RSD model with fixed \( \mu \) (i.e., \( \mu = 3 \)). For this reason we have
focused on the retrieval of \( D_0 \) and \( N_w \) only. Once the
polarimetric variables have been generated, a zero mean

Figure 1 Error histograms of the retrieved RSD parameters. The
upper panels refer to results obtained applying the NN technique. The
middle and lower panels to the \( \beta \) and C-G methods, respectively.

Figure 2 Error histograms of the estimated rain rate. The upper panel
shows the results obtained applying the proposed NN technique. The
middle and lower panels show the corresponding results for the \( \beta \) and
C-G methods, respectively.
algorithms, the corresponding rain rates can be computed using (5). $R_{NN}$, $R_{\beta}$ and $R_{C-G}$ denote the rain rate retrieval algorithms based on the neural network, $\beta$, and C-G methods, respectively (see Figure 2). It has been found that the errors are less than 10 mm h$^{-1}$ in about 98.2, 85.8 and 90.6% of the cases respectively for $R_{NN}$, $R_{\beta}$ and $R_{C-G}$. It is worth mentioning that, while the neural network based technique shows a better performance ($\text{RMSE} \sim 3.7$ mm h$^{-1}$, $r=0.951$), the differences among the other algorithms are small. $R_{\beta}$ provides better results in terms of RMSE ($\sim 8$ mm h$^{-1}$) but a reduced correlation ($r=0.758$), the corresponding values for $R_{C-G}$ being about 11.4 and 0.899.

5 Conclusions

A new neural network algorithm to estimate the raindrop size distribution from S-band dual-polarized radar measurements is proposed in this work. The neural-network algorithm exhibits enhanced features to improve its efficiency, robustness and generalization capability. Numerical simulations, performed using a T-matrix hydrometeor scattering model, have been used to investigate the accuracy of the proposed methodology. The precipitation model has been characterized in terms of shape, raindrop size distribution and orientation. A disdrometer-derived stochastic model has been employed to simulate RSD variability and to construct a test data set. The error analysis, performed in order to evaluate the expected errors of this method, have shown an improvement with respect to other methodologies described in the literature to estimate RSD parameters and, consequently, the corresponding rain rates.

Acknowledgements: This work has been partially funded by European project Risk-AWARE within INTERREG-IIIB-CADSES and by the Italian Ministry of University and Research (MIUR). One of the authors (VC) acknowledges support from the NSF-ERC program (CASA-ERC-0313747). The author A.B. is financially supported by the EU Integrated Project FLOODsite (GOCE-CT-2004-505420). The author R.U. is financially supported by the Netherlands Organization for Scientific Research (NWO) through a Vernieuwingsimpuls/VIDI grant (Project 016.021.003).

References


Bishop, C., 1996: Neural Networks for Pattern Recognition, Clarendon Press.


